

**PCT**ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE  
Bureau international

## DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

<b>(51) Classification internationale des brevets <sup>6</sup> :</b> C12N 15/62, 15/14, 15/12, 15/60, 15/80, 15/81, 1/19, 1/15, A61K 38/38, 38/17, C07K 14/765 // C12N 9/88, 1/21, 5/10	<b>A1</b>	<b>(11) Numéro de publication internationale:</b> WO 95/30759 <b>(43) Date de publication internationale:</b> 16 novembre 1995 (16.11.95)
<b>(21) Numéro de la demande internationale:</b> PCT/FR95/00520 <b>(22) Date de dépôt international:</b> 20 avril 1995 (20.04.95) <b>(30) Données relatives à la priorité:</b> 94/05616 6 mai 1994 (06.05.94) FR <b>(71) Déposant (pour tous les Etats désignés sauf US):</b> RHONE-POULENC RORER S.A. [FR/FR]; 20, avenue Raymond-Aron, F-92160 Antony (FR). <b>(72) Inventeurs; et</b> <b>(75) Inventeurs/Déposants (US seulement):</b> BECQUART, Jérôme [FR/FR]; 8, rue Maublanc, F-75015 Paris (FR). CON-SEILLER, Emmanuel [FR/FR]; 10, rue de Plélo, F-75015 Paris (FR). GUITTON, Jean-Dominique [FR/FR]; 74, rue Dunois, F-75013 Paris (FR). HARDY, Florence [FR/FR]; 6, square Villaret-de-Joyeuse, F-75017 Paris (FR). YEH, Patrice [FR/FR]; 11 bis, rue Lacépède, F-75005 Paris (FR). <b>(74) Mandataire:</b> LE COUPANEC, Pascale; Rhône-Poulenc Rorer S.A., Direction Brevets, 20, avenue Raymond-Aron, F-92160 Antony (FR).	<b>(81) Etats désignés:</b> CA, JP, US, brevet européen (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Publiée</b> <i>Avec rapport de recherche internationale.</i>	
<b>(54) Title:</b> BIOLOGICALLY ACTIVE POLYPEPTIDES INSERTED INTO AN ALBUMIN <b>(54) Titre:</b> POLYPEPTIDES BIOLOGIQUEMENT ACTIFS INSERES DANS UNE ALBUMINE <b>(57) Abstract</b> <p>Biologically active recombinant polypeptides essentially consisting of at least one active portion derived from a natural or artificial biologically active polypeptide and inserted into an albumin or albumin variant, the preparation thereof, and pharmaceutical compositions containing same, are disclosed.</p> <b>(57) Abrégé</b> <p>La présente invention concerne des polypeptides recombinants biologiquement actifs essentiellement composés d'au moins une partie active dérivée d'un polypeptide, naturel ou artificiel, ayant une activité biologique, insérée dans une albumine ou un variant d'albumine, leur préparation et des compositions pharmaceutiques les contenant.</p>		

# **UNIQUEMENT A TITRE D'INFORMATION**

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AT	Autriche	GB	Royaume-Uni	MR	Mauritanie
AU	Australie	GE	Géorgie	MW	Malawi
BB	Barbade	GN	Guinée	NE	Niger
BE	Belgique	GR	Grèce	NL	Pays-Bas
BF	Burkina Faso	HU	Hongrie	NO	Norvège
BG	Bulgarie	IE	Irlande	NZ	Nouvelle-Zélande
BJ	Bénin	IT	Italie	PL	Pologne
BR	Brésil	JP	Japon	PT	Portugal
BY	Bélarus	KE	Kenya	RO	Roumanie
CA	Canada	KG	Kirghizistan	RU	Fédération de Russie
CF	République centrafricaine	KP	République populaire démocratique de Corée	SD	Soudan
CG	Congo	KR	République de Corée	SE	Suède
CH	Suisse	KZ	Kazakhstan	SI	Slovénie
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovaquie
CM	Cameroun	LK	Sri Lanka	SN	Sénégal
CN	Chine	LU	Luxembourg	TD	Tchad
CS	Tchécoslovaquie	LV	Lettonie	TG	Togo
CZ	République tchèque	MC	Monaco	TJ	Tadjikistan
DE	Allemagne	MD	République de Moldova	TT	Trinité-et-Tobago
DK	Danemark	MG	Madagascar	UA	Ukraine
ES	Espagne	ML	Mali	US	Etats-Unis d'Amérique
FI	Finlande	MN	Mongolie	UZ	Ouzbékistan
FR	France			VN	Viet Nam
GA	Gabon				

## Polypeptides biologiquement actifs insérés dans une albumine

La présente invention concerne de nouveaux polypeptides biologiquement  
5 actifs, leur préparation et des compositions pharmaceutiques les contenant.

Quoique possédant une ou plusieurs activités thérapeutiques potentielles, de nombreux polypeptides ne peuvent, malheureusement, pas être exploités pharmaceutiquement. Ceci peut avoir différentes raisons, telles que notamment leur faible stabilité in vivo, leur structure complexe ou fragile, la difficulté de les produire à  
10 une échelle industriellement acceptable, etc... De même, certains polypeptides ne donnent pas les résultats attendus in vivo en raison de problèmes d'administration, de conditionnement, de pharmacocinétique etc...

La présente invention a précisément pour objet de remédier à ces inconvénients.

15 Elle vise notamment l'élaboration de protéines artificielles biologiquement actives permettant une exploitation optimale, sur le plan thérapeutique, des propriétés biologiques de ces polypeptides.

La Demanderesse a ainsi mis en évidence qu'il est possible d'insérer par voie génétique toute structure active, dérivée d'un polypeptide biologiquement actif, dans  
20 une autre structure protéique constituée d'albumine, sans en altérer lesdites propriétés biologiques. De manière inattendue, la sérum-albumine humaine permet de présenter efficacement la structure active à ses sites d'interaction et d'assurer une stabilité plasmatique élevée au polypeptide recombinant de l'invention.

Plus précisément, la présente invention concerne un polypeptide  
25 recombinant comportant au moins une partie active dérivée d'un polypeptide, naturel ou artificiel, biologiquement actif, génétiquement insérée dans une albumine ou un de ses variants ou dérivés.

Par variant de l'albumine, on entend désigner selon la présente invention toute protéine à haute demi-vie plasmatique obtenue par modification, à l'aide des  
30 techniques du génie génétique, d'un gène codant pour un isomorphe donné de la

sérum-albumine humaine, ainsi que toute macromolécule à haute demie-vie plasmatique obtenue par modification in vitro de la protéine codée par de tels gènes. (Par modification, on doit entendre toute mutation, substitution, délétion, addition ou modification de nature génétique et/ou chimique). L'albumine étant très polymorphe,  
5 de nombreux variants naturels ont été identifiés et répertoriés [Weitkamp L.R. et al., Ann. Hum. Genet. 37 (1973) 219].

En ce qui concerne les dérivés d'albumine, il s'agit plus particulièrement de molécules comportant tout ou partie de l'albumine, fusionnée le cas échéant à au moins une séquence polypeptidique provenant d'un gène naturel ou artificiel, elle  
10 même dotée ou non d'une activité biologique.

Dans la suite de la description, les différents types d'albumines explicités ci-dessus sont communément désignés sous le terme albumine.

Au sens de la présente invention, il est entendu par partie active, une partie possédant une activité qui peut soit être directe (traitement des maladies, diagnostic,  
15 recherche biologique, capteurs...), ou indirecte (par exemple utilisable dans la prévention des maladies, dans la conception des vaccins, dans les techniques de l'imagerie médicale etc...).

Les parties actives de polypeptides biologiquement actifs, insérées selon l'invention, présentent de préférence un intérêt thérapeutique.

20 Les polypeptides possédant une activité thérapeutique peuvent être d'origine humaine ou non.

A titre représentatif des polypeptides d'origine non humaine, on peut citer des peptides ou leurs dérivés, possédant des propriétés potentiellement utiles dans les pathologies des compartiments sanguins et interstitiels, tels que l'hirudine, la  
25 trigramine, l'antistatine, les peptides anticoagulant des tiques (TAP), l'ariétine, l'aplagine etc....

Selon un mode privilégié de l'invention, le polypeptide ayant une activité thérapeutique est un polypeptide d'origine humaine ou un variant moléculaire. Par exemple, il peut s'agir de tout ou partie, d'un enzyme, d'un inhibiteur d'enzyme, d'un  
30 antigène, d'un anticorps, d'une hormone, d'un récepteur, d'un facteur intervenant

dans le contrôle de la coagulation, d'un interféron, d'une cytokine [les interleukines, mais aussi leurs variants antagonistes naturels de leur fixation au(x) récepteur(s), les cytokines de type SIS (small induced secreted) et par exemple les protéines inflammatoires des macrophages (les MIPs), etc...], d'un facteur de croissance et/ou de différenciation [et par exemple les facteurs de croissance transformants (les TGFs), les facteurs de différenciation des cellules sanguines (érythropoïétine, M-CSF, G-CSF, GM-CSF etc...), l'insuline et les facteurs de croissance qui lui ressemblent (les IGFs), ou encore les facteurs de perméabilité cellulaire (VPF/VEGF), etc...], d'un facteur impliqué dans la génèse/résorption des tissus osseux (OIF et ostéopontine par exemple), d'un facteur impliqué dans la motilité ou la migration cellulaire [et par exemple le facteur de mobilité autocrine (AMF), le facteur de stimulation de la migration (MSF), ou encore le facteur de dispersion (scatter factor/facteur de croissance des hépatocytes)], d'un facteur bactéricide ou antifongique, d'un facteur chimiotactique [et par exemple le facteur plaquettaire 4 (PF4), ou encore les peptides chemoattractants des monocytes (MCP/MCAF) ou des neutrophiles (NCAF), etc...], d'un facteur cytostatique (et par exemple les protéines qui se fixent aux galactosides), d'une molécule adhésive plasmatique (et par exemple le facteur de von Willebrand, le fibrinogène etc...) ou interstitielle (laminine, ténascine, vitronectine, etc...) ou des matrices extracellulaires, ou encore toute séquence peptidique antagoniste ou agoniste d'interactions moléculaires et/ou intercellulaires impliquées dans les pathologies des compartiments circulatoires et interstitiels et par exemple la formation des thrombus artériels et veineux, les métastases cancéreuses, l'angiogénèse tumorale, le choc inflammatoire, les maladies autoimmunes, les pathologies osseuses et ostéo-articulaires etc...

Bien entendu, la partie active des polypeptides de l'invention peut être constituée par le polypeptide entier biologiquement actif ou par une structure dérivée de celui-ci ou encore correspondre à une séquence peptidique non naturelle, isolée par exemple à partir de banques peptidiques aléatoires. (Pour des raisons de simplification, ces différentes possibilités seront couvertes dans ce qui suit sous la désignation commune "partie active d'un peptide biologiquement actif".) Au sens de la présente invention, on entend par structure dérivée tout polypeptide obtenu par modification et conservant une activité biologique. Par modification, on doit entendre toute mutation, substitution, délétion, addition ou modification de nature génétique et/ou chimique. De tels dérivés peuvent être générés dans des buts différents, tels

que notamment celui d'augmenter l'affinité de la molécule pour ses sites de fixation, celui d'améliorer ses niveaux de production, celui d'augmenter sa résistances aux protéases, celui d'augmenter son efficacité thérapeutique ou encore de réduire ses effets secondaires, ou celui de lui conférer de nouvelles propriétés biologiques. A titre  
5 d'exemple, les polypeptides chimères de l'invention possèdent des propriétés pharmacocinétiques et une activité biologique utilisable pour la prévention ou le traitement des maladies.

Des polypeptides de l'invention, particulièrement avantageux, sont ceux dans lesquels la partie active présente :

- 10 (a) la structure peptidique entière ou,  
(b) un fragment de (a) ou une structure dérivée de (a) par modification structurale (mutation, substitution addition et/ou délétion d'un ou plusieurs résidus) et possédant une activité thérapeutique.

Parmi les structures du type (b), on peut citer plus particulièrement les  
15 molécules dans lesquelles certains sites de N- ou O-glycosylation ont été modifiés ou supprimés, les molécules dans lesquelles un ou plusieurs résidus ont été substitués, ou les molécules dans lesquelles tous les résidus cystéine ont été substitués. On peut citer également des molécules obtenues à partir de (a) par délétion de régions n'intervenant pas ou peu dans l'interaction avec les sites de liaison considérés, ou  
20 exprimant une activité indésirable, et des molécules comportant par rapport à (a) des résidus supplémentaires, tels que par exemple une méthionine N-terminale et/ou un signal de sécrétion et/ou un peptide de jonction.

L'objet de la présente invention est particulièrement avantageux pour les peptides actifs trop petits pour former un domaine structural et/ou ne possédant pas  
25 une bonne stabilité in vivo et/ou une bonne biodisponibilité. L'insertion proposée selon l'invention, permet de les associer à un ou des domaine(s) pré-existant(s) de l'albumine et de bénéficier ainsi de la biodisponibilité et de la stabilité in vivo de celle-ci.

De manière générale, la taille des parties actives insérées dans l'albumine  
30 varie entre trois à vingt cinq résidus d'acides aminés. Toutefois, des séquences allant de 1 résidu à 100 résidus peuvent également être utilisées.

L'insertion d'une partie active de peptide dans la séquence peptidique de l'albumine est réalisée selon l'invention de manière à satisfaire aux deux conditions suivantes:

Il doit être préservé à ladite partie active, insérée au sein de l'albumine, une  
5 accessibilité suffisante afin de lui conserver intacte son activité biologique. Par ailleurs, la structure de l'albumine ne doit pas non plus subir une déstabilisation trop importante qui serait bien entendu préjudiciable au polypeptide recombinant dit la chimère.

Les sites d'insertion sont préférentiellement sélectionnés au sein de  
10 l'albumine en respectant les précédents impératifs.

Selon la structure cristalline publiée par He et Carter (Nature 1992, 358, 209-215) l'albumine est formée de la répétition de 3 domaines comprenant chacun deux sous-domaines et elle contient plus de 67% d'hélices alpha. Chacun des domaines est superposable aux autres et est formé de 10 hélices notées de h1 à h10. Le sous-  
15 domaine A comporte les hélices h1 à h6 et le sous-domaine B, les hélices h7 à h10. Chaque sous domaine est formé par un motif commun: h1, h2, h3, h4 pour le domaine A et h7, h8, h9 et h10 pour le domaine B. Les petites hélices h5 et h6 supplémentaires sont liées par un pont disulfure au sous domaine A. La figure 1 rend compte de manière schématique de la structure de l'albumine humaine.

20 Les sites d'insertion sont de préférence localisés dans les régions de l'albumine présumées former des régions exposées à la surface de la molécule, ces régions étant préférentiellement des boucles.

A titre de sites d'insertion convenant tout particulièrement à l'invention, on peut mentionner trois régions du premier domaine :

- 25 - la région 5 qui s'étend du résidu 57 à 62 et qui correspond à une boucle reliant les deux hélices h3 et h4 ;  
- la région 8 comprenant les résidus 103 à 120 et correspondant à la zone inter sous-domaines.  
- la région 13 comprise entre les résidus 178 et 200 et qui correspond à une hélice.

30 Les hélices h2 et h3 du domaine III délimitent également, du résidu 419 au résidu 430, une autre région d'insertion envisageable selon l'invention.

Une partie active de peptide biologiquement actif peut être insérée selon trois modes différents dans la séquence peptidique de l'albumine.

- Il peut s'agir d'une insertion stricte consistant en une simple addition de la séquence du peptide d'intérêt dans la séquence d'origine de l'albumine qui est conservée dans sa totalité.
- L'insertion peut correspondre à une substitution d'une portion de la séquence peptidique de l'albumine par la séquence peptidique correspondant à la partie active du peptide d'intérêt.
- Enfin, il peut s'agir d'une insertion combinant une addition d'une partie de la séquence peptidique active et une substitution d'une portion de la séquence peptidique de l'albumine par le reste de la partie active du peptide actif.

La figure 2 rend compte d'une manière schématique de ces différents modes d'insertion.

Bien entendu, la partie active d'un peptide biologiquement actif peut être répétée plusieurs fois dans la chimère au même endroit et/ou dans des régions différentes de l'albumine. De même, il est également possible d'insérer selon l'invention des parties actives différentes, soit issues d'un même peptide ou de peptides différents.

Enfin, une partie active des peptides selon l'invention peut être insérée soit strictement au sein de l'albumine soit entourée de séquences de jonction.

En ce qui concerne ces séquences de jonction, il peut notamment s'agir de séquences peptidiques riches en résidus glycine et/ou en résidus sérine et/ou en résidus thréonine et/ou tout résidu d'acide aminé décrit comme fréquemment rencontré dans les zones de flexibilité dans les protéines.

Les polypeptides recombinants de l'invention s'avèrent tout particulièrement avantageux.

Ils permettent de maintenir dans l'organisme une activité biologique donnée pendant un temps prolongé. Il s'avère ainsi possible de réduire les doses administrées et, dans certains cas, de potentialiser l'effet thérapeutique, par exemple en réduisant les effets secondaires consécutifs à une administration plus importante. Ils permettent avantageusement de générer et d'utiliser des structures dérivées de polypeptides biologiquement actifs très petites et donc très spécifiques d'un effet recherché. Ces polypeptides recombinants possèdent par ailleurs une répartition particulièrement avantageuse dans l'organisme, ce qui modifie leurs propriétés pharmacocinétiques et favorise le développement de leur activité biologique et leur



utilisation. Ils présentent également l'avantage d'être faiblement ou non-immunogéniques pour l'organisme dans lequel ils sont utilisés. Finalement, les polypeptides de l'invention peuvent être exprimés (et préférentiellement sécrétés) par des organismes recombinants, à des niveaux permettant leur exploitation industrielle.

5           Un autre objet de l'invention concerne un procédé de préparation des molécules chimères décrites ci-avant. Plus précisément, ce procédé consiste à faire exprimer par un hôte cellulaire eucaryote ou procaryote une séquence nucléotidique codant pour une partie active d'un polypeptide désiré, puis à récolter le polypeptide produit.

10           Parmi les hôtes eucaryotes utilisables dans le cadre de la présente invention, on peut citer les cellules animales, les levures, ou les champignons. En particulier, s'agissant de levures, on peut citer les levures du genre Saccharomyces, Kluyveromyces, Pichia, Schwanniomyces, ou Hansenula. S'agissant de cellules animales, on peut citer les cellules COS, CHO, Cl27, etc... Parmi les champignons  
15 susceptibles d'être utilisés dans la présente invention, on peut citer plus particulièrement Aspergillus ssp. ou Trichoderma ssp. Comme hôtes procaryotes, on préfère utiliser les bactéries telles que Escherichia coli, ou appartenant aux genres Corynebacterium, Bacillus, ou Streptomyces.

20           Les séquences nucléotidiques utilisables dans le cadre de la présente invention peuvent être préparées de différentes manières. Généralement, elles sont obtenues en assemblant en phase de lecture les séquences codant pour chacune des parties fonctionnelles du polypeptide. Celles-ci peuvent être isolées par les techniques de l'homme de l'art, et par exemple directement à partir des ARN messagers (ARNm) cellulaires, ou par reclonage à partir d'une banque d'ADN  
25 complémentaire (ADNc), ou encore il peut s'agir de séquences nucléotidiques totalement synthétiques. Il est entendu de plus que les séquences nucléotidiques peuvent également être ultérieurement modifiées, par exemple par les techniques du génie génétique, pour obtenir des dérivés ou des variants desdites séquences.

30           L'insertion de cette séquence nucléotidique, codant pour la partie active du polypeptide, peut être réalisée directement ou non, selon la région choisie pour site d'insertion, dans le gène codant pour l'albumine.

La région sélectionnée peut, en effet, ne pas comporter de site de restriction adéquat à la réalisation de ladite insertion. Dans cette hypothèse, il peut s'avérer intéressant, préalablement à l'insertion, d'introduire un ou plusieurs sites de restriction

uniques. La création de sites de restriction, au niveau de la région choisie pour site d'insertion, se fait de préférence par mutagénèse dirigée selon des techniques classiques. Toutefois on peut également insérer la séquence nucléotidique correspondant à "la partie active du polypeptide biologiquement actif" directement par  
5 mutagénèse dirigée sans que l'insertion ne fasse apparaître de sites de restriction particuliers.

Dans le cas particulier de la région 5 du gène codant pour l'albumine, la présence du site de restriction PVuII autorise l'insertion directe de la séquence nucléotidique. Il se révèle donc particulièrement utile comme site de clonage d'une  
10 partie active d'un polypeptide que l'on souhaite insérer en phase traductionnelle dans la séquence de l'albumine au niveau du 57<sup>ème</sup> résidu.

Dans ce mode d'insertion, mettant à profit un site de restriction unique déjà existant dans la séquence d'origine de l'albumine, la ligature de la séquence codant pour le peptide actif avec le fragment de restriction, correspondant à la totalité du  
15 gène codant pour l'albumine, génère une séquence nucléotidique comportant un gène hybride codant pour une protéine chimère du type SAH insertion stricte.

Dans le cas particulier de la zone 419 à 430, la présence de 2 sites uniques d'insertion, (HincII et ArvII) permet d'insérer et/ou substituer les peptides d'intérêt biologique à la séquence de l'albumine.

20 En ce qui concerne plus particulièrement les régions 8 et 13, l'insertion de ladite séquence est favorisée si l'on y crée préalablement un ou des sites de restriction manipulables. Bien entendu, les sites de restriction à créer au niveau de la séquence sont choisis en tenant compte de la nature des sites de restriction déjà existants. Ils doivent conduire à une insertion sélective.

25 Dans ce second mode de réalisation, l'emploi de deux sites de restriction uniques permet de construire des gènes codant pour des chimères ayant le peptide actif en insertion et/ou en substitution. L'insertion du peptide se fait par remplacement d'un fragment borné de deux sites de restriction uniques, dans l'ADN complémentaire de l'albumine, insérés par mutagénèse dirigée. Ces deux sites de restriction uniques  
30 pourront être, Mst I et Kpn I, dans la région 8 et Sst I et Xho I dans la région 13. La création de ces sites de restriction peut ou non modifier la séquence polypeptidique de l'albumine humaine. Le clonage subséquent du peptide actif en phase codante dans ce gène de l'albumine peut être exclusivement la séquence codante du peptide

ou peut correspondre à un panachage de la séquence codante du peptide et de la séquence codante du fragment délété de l'albumine.

La présente invention vise également la protection des variants de  
5 séquences nucléotidiques codant pour l'albumine correspondante c'est à dire intégrant au moins un site de restriction unique non naturel c'est à dire non présent dans la séquence d'origine.

Avantageusement, la création de ces sites de restriction peut servir par la suite à l'insertion d'un ou plusieurs parties actives de polypeptide(s) biologiquement  
10 actif(s) dans la protéine mature.

Plus préférentiellement, dans le procédé de l'invention, la séquence nucléotidique fait partie d'une cassette d'expression comprenant une région d'initiation de la transcription (région promoteur) permettant, dans les cellules hôtes, l'expression  
15 de la séquence nucléotidique placée sous son contrôle et codant pour les polypeptides de l'invention. Cette région peut provenir de régions promoteurs de gènes fortement exprimés dans la cellule hôte utilisée, l'expression étant constitutive ou régulable. S'agissant de levures, il peut s'agir du promoteur du gène de la phosphoglycérate kinase (PGK), de la glyceraldéhyde-3-phosphate déshydrogénase (GPD), de la lactase (LAC4), des énalases (ENO), des alcools deshydrogénases (ADH), etc... S'agissant de bactéries, il peut s'agir du promoteur des gènes droit ou  
20 gauche du bactériophage lambda (P<sub>L</sub>, P<sub>R</sub>), ou encore des promoteurs des gènes des opérons tryptophane (P<sub>trp</sub>) ou lactose (P<sub>lac</sub>). En outre, cette région de contrôle peut être modifiée, par exemple par mutagénèse in vitro, par introduction d'éléments  
25 additionnels de contrôle ou de séquences synthétiques, ou par des délétions ou des substitutions des éléments originels de contrôle. La cassette d'expression peut également comprendre une région de terminaison de la transcription fonctionnelle dans l'hôte envisagé, positionnée immédiatement en aval de la séquence nucléotidique codant pour un polypeptide de l'invention.

30 Dans un mode préféré, les polypeptides de l'invention résultent de l'expression dans un hôte eucaryote ou procaryote d'une séquence nucléotidique et de la sécrétion du produit d'expression de ladite séquence dans le milieu de culture. Il est en effet particulièrement avantageux de pouvoir obtenir par voie recombinante des molécules directement dans le milieu de culture. Dans ce cas, la séquence

nucléotidique codant pour un polypeptide de l'invention est précédée d'une séquence "leader" (ou séquence signal) dirigeant le polypeptide naissant dans les voies de sécrétion de l'hôte utilisé. Cette séquence "leader" peut être la séquence signal naturelle du polypeptide biologiquement actif dans le cas où celui-ci est une protéine  
5 naturellement sécrétée, ou celle de la structure stabilisatrice, mais il peut également s'agir de toute autre séquence "leader" fonctionnelle, ou d'une séquence "leader" artificielle. Le choix de l'une ou l'autre de ces séquences est notamment guidé par l'hôte utilisé. Des exemples de séquences signal fonctionnelles incluent celles des gènes des phéromones sexuelles ou des toxines "killer" de levures.

10 En plus de la cassette d'expression, un ou plusieurs marqueurs permettant de sélectionner l'hôte recombiné peuvent être additionnés, tels que par exemple le gène URA3 de la levure S. cerevisiae, ou des gènes conférant la résistance à des antibiotiques comme la généticine (G418) ou à tout autre composé toxique comme certains ions métalliques.

15 L'ensemble constitué par la cassette d'expression et par le marqueur de sélection peut être introduit directement dans les cellules hôtes considérées, soit inséré préalablement dans un vecteur autorépliquatif fonctionnel. Dans le premier cas, des séquences homologues à des régions présentes dans le génôme des cellules hôtes sont préférentiellement additionnées à cet ensemble; lesdites séquences étant  
20 alors positionnées de chaque côté de la cassette d'expression et du gène de sélection de façon à augmenter la fréquence d'intégration de l'ensemble dans le génôme de l'hôte en ciblant l'intégration des séquences par recombinaison homologue. Dans le cas où la cassette d'expression est insérée dans un système répliquatif, un système de réplication préféré pour les levures du genre Kluyveromyces  
25 est dérivé du plasmide pKD1 initialement isolé de K. drosophilum; un système préféré de réplication pour les levures du genre Saccharomyces est dérivé du plasmide 2 $\mu$  de S. cerevisiae. De plus, ce plasmide d'expression peut contenir tout ou partie desdits systèmes de réplication, ou peut combiner des éléments dérivés du plasmide pKD1 aussi bien que du plasmide 2 $\mu$ .

30 En outre, les plasmides d'expression peuvent être des vecteurs navettes entre un hôte bactérien tel que Escherichia coli et la cellule hôte choisie. Dans ce cas, une origine de réplication et un marqueur de sélection fonctionnant dans l'hôte bactérien sont requises. Il est également possible de positionner des sites de

restriction entourant les séquences bactériennes et uniques sur le vecteur d'expression: ceci permet de supprimer ces séquences par coupure et religature in vitro du vecteur tronqué avant transformation des cellules hôtes, ce qui peut résulter en une augmentation du nombre de copies et en une stabilité accrue des plasmides d'expression dans lesdits hôtes. Par exemple, de tels sites de restriction peuvent correspondre aux séquences telles que 5'-GGCCNNNNNGGCC-3' (SfiI) (SEQ ID N°1) ou 5'-GCGGCCGC-3' (NotI) (SEQ ID N°2) dans la mesure où ces sites sont extrêmement rares et généralement absents d'un vecteur d'expression.

Après construction de tels vecteurs ou cassette d'expression, ceux-ci sont introduits dans les cellules hôtes retenues selon les techniques classiques décrites dans la littérature. A cet égard, toute méthode permettant d'introduire un ADN étranger dans une cellule peut être utilisée. Il peut s'agir notamment de transformation, électroporation, conjugaison, ou toute autre technique connue de l'homme de l'art. A titre d'exemple pour les hôtes de type levure, les différentes souches de Kluyveromyces utilisées ont été transformées en traitant les cellules entières en présence d'acétate de lithium et de polyéthylène glycol, selon la technique décrite par Ito et al. [J. Bacteriol. 153 (1983) 163]. La technique de transformation décrite par Durrens et al. [Curr. Genet. 18 (1990) 7] utilisant l'éthylène glycol et le diméthylsulfoxyde a également été utilisée. Il est aussi possible de transformer les levures par électroporation, selon la méthode décrite par Karube et al. [FEBS Letters 182 (1985) 90]. Un protocole alternatif est également décrit en détail dans les exemples qui suivent.

Après sélection des cellules transformées, les cellules exprimant lesdits polypeptides sont inoculées et la récupération desdits polypeptides peut être faite, soit au cours de la croissance cellulaire pour les procédés "en continu", soit en fin de croissance pour les cultures "en lots" ("batch"). Les polypeptides qui font l'objet de la présente invention sont ensuite purifiés à partir du surnageant de culture en vue de leur caractérisation moléculaire, pharmacocinétique et biologique.

Un système d'expression préféré des polypeptides de l'invention consiste en l'utilisation des levures du genre Kluyveromyces comme cellule hôte, transformées par certains vecteurs dérivés du réplicon extrachromosomique pKD1 initialement isolé chez K. marxianus var. drosophilaram. Ces levures, et en particulier K. lactis et K. fragilis sont généralement capables de répliquer lesdits vecteurs de façon stable et possèdent en outre l'avantage d'être incluses dans la liste des organismes G.R.A.S.

("Generally Recognized As Safe"). Des levures privilégiées sont préférentiellement des souches industrielles du genre Kluyveromyces capables de répliquer de façon stable lesdits plasmides dérivés du plasmide pKD1 et dans lesquels a été inséré un marqueur de sélection ainsi qu'une cassette d'expression permettant la sécrétion à des niveaux élevés des polypeptides de l'invention.

La présente invention concerne également les séquences nucléotidiques codant pour les polypeptides chimères décrits ci-avant, ainsi que les cellules recombinantes, eucaryotes ou procaryotes, comprenant de telles séquences.

La présente invention concerne aussi l'application à titre de médicament des polypeptides selon la présente invention. Plus particulièrement, l'invention a pour objet toute composition pharmaceutique comprenant un ou plusieurs polypeptides ou séquences nucléotidiques tels que décrits ci-avant. Ces compositions pharmaceutiques peuvent se présenter sous forme de diverses formulations. Il peut notamment s'agir de nanoparticules à la surface desquelles sont présents des polypeptides selon l'invention. Ce type de formulation est plus particulièrement utilisé pour effectuer des ciblage dirigés en principe actif. Bien entendu, les séquences nucléotidiques peuvent être utilisées en thérapie génique.

La présente invention sera plus complètement décrite à l'aide des exemples qui suivent, qui doivent être considérés comme illustratifs et non limitatifs.

20

### LISTE DES FIGURES

Les représentations des plasmides indiquées dans les Figures suivantes ne sont pas tracées à l'échelle et seuls les sites de restriction importants pour la compréhension des clonages réalisés ont été indiqués.

25 Figure 1: Représentation schématique du domaine I de l'albumine avec localisation de sites d'insertion selon l'invention. \* signale la localisation du site d'insertion h2-h3 dans le domaine III. Les chiffres 5, 8 et 13 identifient les zones d'insertion correspondantes.

30 Figure 2 : Représentation schématique de différents modes d'insertion d'un peptide actif dans la structure de l'albumine.

Figure 3: Plasmide pYG105.

Figure 4: Modification de la région 5 de l'albumine suite à l'insertion de la séquence codant pour IEGR telle que décrite dans l'exemple 8.1. Les modifications figurent en caractères gras. L'emplacement des sites de restriction modifiés ou apportés sont indiqués par un trait horizontal et la position de coupure des enzymes correspondant par un trait vertical.

Figure 5: Stratégie de clonage pour l'insertion de IEGR dans la région 5 (exemple 8.1).

Figure 6: Stratégie de clonage pour l'insertion d'un peptide actif dans la région 13 de l'albumine.

## **EXEMPLES**

### **TECHNIQUES GENERALES DE CLONAGE**

Les méthodes classiquement utilisées en biologie moléculaire telles que les extractions préparatives d'ADN plasmidique, la centrifugation d'ADN plasmidique en gradient de chlorure de césium, l'électrophorèse sur gels d'agarose ou d'acrylamide, la purification de fragments d'ADN par électroélution, les extraction de protéines au phénol ou au phénol-chloroforme, la précipitation d'ADN en milieu salin par de l'éthanol ou de l'isopropanol, la transformation dans *Escherichia coli* etc ... sont bien connues de l'homme de métier et sont abondamment décrites dans la littérature [Maniatis T. et al., "Molecular Cloning, a Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982 ; Ausubel F.M. et al. (eds), "Current Protocols in Molecular Biology", John Wiley & Sons, New York, 1987].

Les enzymes de restriction ont été fournies par New England Biolabs (Biolabs), Bethesda Research Laboratories (BRL) ou Amersham et sont utilisées selon les recommandations des fournisseurs.

Les plasmides de type pBR322, pUC et les phages de la série M13 sont d'origine commerciale (Bethesda Research Laboratories).

Pour les ligatures, les fragments d'ADN sont séparés selon leur taille par électrophorèse en gels d'agarose ou d'acrylamide, extraits au phénol ou par un

mélange phénol/chloroforme, précipités à l'éthanol puis incubés en présence de l'ADN ligase du phage T4 (Biolabs) selon les recommandations du fournisseur.

Le remplissage des extrémités 5' proéminentes est effectué par le fragment de Klenow de l'ADN Polymérase I d'E. coli (Biolabs) selon les spécifications du fournisseur. La destruction des extrémités 3' proéminentes est effectuée en présence  
5 de l'ADN Polymérase du phage T4 (Biolabs) utilisée selon les recommandations du fabricant. La destruction des extrémités 5' proéminentes est effectuée par un traitement ménagé par la nucléase S1.

La mutagénèse dirigée in vitro par oligodéoxynucléotides synthétiques est effectuée selon la méthode développée par Taylor et al. [Nucleic Acids Res. 13  
10 (1985) 8749-8764] en utilisant le kit distribué par Amersham. Cette technique est notamment mise en oeuvre, dans le cadre de la présente invention, pour créer des sites de restriction uniques en vu d'une insertion subséquente.

L'amplification enzymatique de fragments d'ADN par la technique dite de PCR  
15 [Polymérase-catalyzed Chain Reaction, Saiki R.K. et al., Science 230 (1985) 1350-1354; Mullis K.B. et Faloona F.A., Meth. Enzym. 155 (1987) 335-350] est effectuée en utilisant un "DNA thermal cycler" (Perkin Elmer Cetus) selon les spécifications du fabricant.

La vérification des séquences nucléotidiques est effectuée par la méthode  
20 développée par Sanger et al. [Proc. Natl. Acad. Sci. USA, 74 (1977) 5463-5467] en utilisant le kit distribué par Amersham.

Les transformations de K. lactis avec l'ADN des plasmides d'expression des protéines de la présente invention sont effectuées par toute technique connue de l'homme de l'art, et dont un exemple est donné dans le texte.

25 Sauf indication contraire, les souches bactériennes utilisées sont E. coli MC1060 (lacIPOZYA, X74, galU, galK, strA<sup>r</sup>), ou E. coli TG1 (lac, proA,B, supE, thi, hsdD5 / F'traD36, proA<sup>+</sup>B<sup>+</sup>, lacI<sup>q</sup>, lacZ, M15).

Les souches de levures utilisées appartiennent aux levures bourgeonnantes et plus particulièrement aux levures du genre Kluyveromyces. Les souche K. lactis  
30 MW98-8C (a, uraA, arg, lys, K<sup>+</sup>, pKD1<sup>o</sup>) et K. lactis CBS 293.91 ont été particulièrement utilisées; un échantillon de la souche MW98-8C a été déposé le 16 Septembre 1988 au Centraalbureau voor Schimmelkulturen (CBS) à Baarn (Pays Bas) où il a été enregistré sous le numéro CBS 579.88.



Une souche bactérienne (E. coli) transformée avec le plasmide pET-8c52K a été déposée le 17 Avril 1990 auprès de l'American Type Culture Collection sous le numéro ATCC 68306.

Les souches de levures transformées par les plasmides d'expression codant  
5 pour les protéines de la présente invention sont cultivées en erlenmeyers ou en fermenteurs pilotes de 2l (SETRIC, France) à 28°C en milieu riche (YPD: 1% yeast extract, 2% Bactopeptone, 2% glucose; ou YPL: 1% yeast extract, 2% Bactopeptone, 2% lactose) sous agitation constante.

10

**EXEMPLE 1:**

**PROTOCOLE D'INSERTION STRICTE DE PEPTIDE EN UTILISANT UN SITE DE RESTRICTION UNIQUE, PRESENT DANS LE GENE DE LA SAH.**

15 De par son unicité dans le gène HSA et le vecteur associé, le site PvuII localisé naturellement dans la séquence codante, est particulièrement utile comme site de clonage d'un peptide biologiquement actif que l'on désire insérer en phase traductionnelle dans la SAH au niveau du 58<sup>ème</sup> résidu. Dans un mode de réalisation particulier, il est utile d'employer des peptides de p résidus dont la  
20 séquence codante est [3xN]<sub>p</sub>. Dans ce cas, les oligonucléotides synthétisés sont du type 5'-NN [3xN]<sub>p</sub>N-3' et son brin complémentaire. La ligature de ce fragment avec le fragment de restriction HindIII-HindIII Δ PvuII, correspondant à la totalité du gène codant pour la SAH, génère un fragment de restriction HindIII-HindIII comportant un gène hybride codant pour une protéine chimère du type SAH insertion stricte. Dans  
25 un autre mode de réalisation, le peptide peut être répété plusieurs fois dans la chimère.

**EXEMPLE 2:**

30 **PROTOCOLE D'INSERTION TOTALE OU PARTIELLE DE PEPTIDE EN UTILISANT DEUX SITES DE RESTRICTION UNIQUES ET NATURELS.**

Dans un mode de réalisation particulier, l'utilisation de deux sites de restriction uniques permet de construire des gènes codant pour des chimères ayant le peptide actif en insertion stricte ou en substitution ou en insertion partielle.

Par exemple, l'existence des sites uniques HincII et AvrII dans la séquence codante de SAH et dans le vecteur permet de générer un fragment HindIII-HindIII Δ HincII-AvrII. L'élimination du fragment nucléotidique HincII-AvrII correspond à la délétion du fragment peptidique T(420) - N(429). L'utilisation d'un oligonucléotide  
5 complémentaire approprié permet de cloner un peptide en phase codante dans le gène de la SAH. Ce fragment de restriction peut être exclusivement la séquence codante complémentaire du peptide ou peut correspondre à un panachage de la séquence codante du peptide actif et de la séquence codante du fragment T(420) - N(429) de la SAH. Le peptide actif peut être présent plusieurs fois dans la chimère.

10

### EXEMPLE 3

#### PLASMIDES D'EXPRESSION

Les protéines chimères des exemples précédents peuvent être exprimées dans les levures à partir de promoteurs fonctionnels, régulables ou constitutifs, tels  
15 que, par exemple, ceux présents dans les plasmides pYG105 (promoteur LAC4 de Kluyveromyces lactis), pYG106 (promoteur PGK de Saccharomyces cerevisiae), pYG536 (promoteur PHO5 de S.cerevisiae), ou des promoteur hybrides tels que ceux décrits dans la demande de brevet EP 361 991. Les plasmides pYG105 et pYG106 sont ici particulièrement utiles car ils permettent l'expression des gènes codés par les  
20 fragments de restriction HindIII tel que décrits dans les exemples précédents et clonés dans le site HindIII et dans l'orientation productive (définie comme l'orientation qui place la région "prépro" de l'albumine de façon proximale par rapport au promoteur de transcription), à partir de promoteurs fonctionnels chez K.lactis, régulables (pYG105) ou constitutifs (pYG106). Le plasmide pYG105 correspond au plasmide pKan707  
25 décrit dans la demande de brevet EP 361 991 dans lequel le site de restriction HindIII unique et localisé dans le gène de résistance à la généticine (G418) a été détruit par mutagenèse dirigée tout en conservant une protéine inchangée (oligodeoxynucleotide 5'-GAAATGCATAAGCTCTTGCCATTCTCACCG-3' (SEQ ID N°3). Le fragment Sall-SacI codant pour le gène URA3 du plasmide muté a été ensuite remplacé par un  
30 fragment de restriction Sall-SacI comportant une cassette d'expression constituée du promoteur LAC4 de K. lactis (sous la forme d'un fragment Sall-HindIII) et du terminateur du gène PGK de S. cerevisiae (sous la forme d'un fragment HindIII-SacI). Le plasmide pYG105 est mitotiquement très stable chez les levures Kluyveromyces. Il

est représenté en figure 3. Les plasmides pYG105 et pYG106 ne diffèrent entre eux que par la nature du promoteur de transcription encodé par le fragment Sall-HindIII.

**EXEMPLE 4:****TRANSFORMATION DES LEVURES**

La transformation des levures appartenant au genre *Kluyveromyces*, et en particulier les souches MW98-8C et CBS 293.91 de *K. lactis*, s'effectue par exemple par la technique de traitement des cellules entières par de l'acétate de lithium [Ito H. et al., J. Bacteriol. 153 (1983) 163-168], adaptée comme suit. La croissance des cellules se fait à 28°C dans 50 ml de milieu YPD, avec agitation et jusqu'à une densité optique à 600 nm (DO<sub>600</sub>) comprise entre 0,6 et 0,8; les cellules sont récoltées par centrifugation à faible vitesse, lavées dans une solution stérile de TE (10 mM Tris HCl pH 7,4; 1 mM EDTA), resuspendues dans 3-4 ml d'acétate lithium (0,1 M dans du TE) pour obtenir une densité cellulaire d'environ  $2 \times 10^8$  cellules/ml, puis incubées à 30°C pendant 1 heure sous agitation modérée. Des aliquotes de 0,1 ml de la suspension résultante de cellules compétentes sont incubés à 30°C pendant 1 heure en présence d'ADN et à une concentration finale de 35% de polyéthylène glycol (PEG<sub>4000</sub>, Sigma). Après un choc thermique de 5 minutes à 42°C, les cellules sont lavées 2 fois, resuspendues dans 0,2 ml d'eau stérile et incubées 16 heures à 28°C dans 2 ml de milieu YPD pour permettre l'expression phénotypique du gène de résistance au G418 exprimé sous contrôle du promoteur P<sub>K1</sub> (cf. EP 361 991); 200 µl de la suspension cellulaire sont ensuite étalés sur boîtes YPD sélectives (G418, 200 µg/ml). Les boîtes sont mises à incuber à 28°C et les transformants apparaissent après 2 à 3 jours de croissance cellulaire.

**EXEMPLE 5:****SECRETION DES CHIMERES**

Après sélection sur milieu riche supplémenté en G418 les clones recombinants sont testés pour leur capacité à sécréter la forme mature des protéines chimères. Quelques clones correspondant à la souche CBS 293.91 ou MW98-8C transformée par les plasmides d'expression des chimères entre la SAH et la partie biologiquement active sont mis à incuber en milieu YPD ou YPL à 28°C. Les sumageants cellulaires sont récupérés par centrifugation quand les cellules atteignent la phase stationnaire de croissance, éventuellement concentrés 10 fois par précipitation pendant 30 minutes à -20°C dans une concentration finale de 60% d'éthanol, puis testés après électrophorèse en gel SDS-PAGE (), soit directement par coloration du gel par du

bleu de coomassie, soit après immunoblot en utilisant des anticorps primaires dirigés contre la partie biologiquement active ou un sérum polyclonal de lapin dirigé contre la SAH. Lors des expériences de détection immunologique, le filtre de nitrocellulose est d'abord incubé en présence des anticorps primaires spécifiques, lavé plusieurs fois, 5 incubé en présence d'anticorps de chèvre dirigés contre les anticorps primaires, puis incubé en présence d'un complexe avidine-péroxydase en utilisant le "kit ABC" distribué par Vectastain (Biosys S.A., Compiègne, France). La réaction immunologique est ensuite révélée par addition de diamino-3,3' benzidine tetrachlorhydrate (Prolabo) en présence d'eau oxygénée, selon les recommandations 10 du fabricant.

#### **EXEMPLE 6:**

##### **PURIFICATION DES CHIMERES**

15

Les chimères présentes dans les surnageants de culture correspondant à la souche CBS 293.91 transformée, sont caractérisées dans un premier temps à l'aide d'anticorps spécifiques de la partie SAH. Il peut être souhaitable de purifier certaines de ces chimères. La culture est alors centrifugée (10000g, 30 min), le surnageant est 20 passé à travers un filtre de 0,22 mm (Millipore), puis concentré par ultrafiltration (Amicon) en utilisant une membrane dont le seuil de discrimination se situe à 30 kDa. Le concentrat obtenu est alors dialysé contre une solution de Tris HCl (50 mM pH 8) puis purifié sur colonne. Par exemple, le concentrat correspondant au surnageant de culture de la souche CBS 293.91 transformée est purifié par chromatographie 25 d'affinité sur Bleu-Trisacryl (IBF) et les échantillons sont ensuite dialysés contre de l'eau. Une purification par tamis moléculaire peut être ensuite réalisée. Dans ce cas, une colonne Superose 12 (Pharmacia) est préalablement équilibrée dans du tampon 20mM NaH<sub>2</sub>PO<sub>4</sub> 100mM NaCl pH 7.0 et les échantillons issus de la chromatographie d'affinité sont chargés sur la colonne, récoltés et caractérisés. Une purification par 30 chromatographie d'échange d'ions peut également être utilisée. Dans certains cas, le concentrat obtenu après ultrafiltration est dialysé contre une solution de Tris HCl (50 mM pH 8), puis déposé par fractions de 20 ml sur une colonne (5 ml) échangeuse de cations (S Fast Flow, Pharmacia) équilibrée dans le même tampon. La colonne est alors lavée plusieurs fois par la solution de Tris HCl (50 mM pH 8) et la protéine

chimère est alors éluée de la colonne par un gradient (0 à 1 M) de NaCl. Les fractions contenant la protéine chimère sont alors réunies et dialysées contre une solution de Tris HCl 50 mM (pH 8) puis redéposées sur colonne S Fast Flow. Après élution de la colonne, les fractions contenant la protéine sont réunies, dialysées  
5 contre de l'eau et lyophilisées avant caractérisation.

#### **EXEMPLE 7:**

#### **INSERTION DU PEPTIDE 11 DANS LA SEQUENCE DE L'ALBUMINE**

10

Le peptide 11 a été décrit en tant qu'épitope de la tryptophane synthase (Larvor et al. Mol. Immunol. (1991) 28, 523-531).

#### **1. Insertion stricte dans une région ayant naturellement des sites de restriction 15 uniques**

Le peptide 11 a la séquence peptidique suivante dans le sens N a C terminal: HGRVGIYFGMK (SEQ ID N°20). La stratégie d'insertion dans la zone HincII-AvrII a consisté à faire une ligature entre un fragment nucléotidique synthétisé codant pour le  
20 peptide 11 et tel que le cadre de lecture pour la séquence codante de la SAH soit respecté et que la nature de la séquence codante de la SAH soit celle d'origine. Dans un premier temps nous avons donc synthétisé les deux oligonucléotides suivants : 5'-CCATGGTAGAGTAGGTATCTATTTCCGGTATGAAAACCTCCAACCTCTGTAGAGGTCTCGAGAAAT-3' (SEQ ID N°4) et 5'-  
25 CTAGATTTCTCGAGACCTCTACAAGATGTGGAGTTTTCATACCGAAATAGATACCTACTCTACCATGG-3' (SEQ ID N°5). Ces deux oligonucléotides ont été hybridés puis ligaturés au fragment HindIII-HindIII  $\Delta$  HincII-AvrII, générant ainsi la totalité du gène de la SAH comportant l'insertion totale du peptide 11 entre S(419) et T(420), immédiatement précédée de la région d'exportation "prépro" de la SAH. Ce fragment est  
30 cloné dans l'orientation productive et dans le site HindIII du plasmide pYG105.

## 2. Insertion par substitution et addition dans une région ayant naturellement des sites de restriction uniques

Dans un autre mode de réalisation, les oligonucléotides synthétisés ont été les suivants : 5'-CCATGGTAGAGTAGGTATCTATTTTCGGTATGAAA-3' (SEQ ID N°6) et 5'-CTAGTTTCATACCGAAATAGATACCTACTTCTTACCATGG-3' (SEQ ID N°7). Ces deux oligonucléotides sont hybridés puis ligaturés au fragment HindIII-HindIII Δ HincII-AvrII, générant ainsi le gène d'une chimère amputé des résidus T(420) à N(429) et substitués par la séquence du peptide 11. Ce fragment est cloné dans l'orientation productive et dans le site HindIII du plasmide pYG105.

## 3. Insertion stricte dans une région ne possédant pas de sites de restriction uniques

Dans un autre mode de réalisation, deux sites de restriction uniques, Sst I et Xho 1, ont été créés par mutagenèse dirigée. Le même type de clonage directionnel a été réalisé. Dans le cas de l'insertion totale du peptide 11 entre les résidus A(191) et S (192), les deux oligonucléotides suivants ont été synthétisés : 5'-ACGGGATGAAGGGAAGGCCCATGGTAGAGTAGGTATCTATTTTCGGTATGAAA-3' (SEQ ID N°8) et 5'-TCGATTTTCATTACCGAAATAGATACCTACTCTACCATGGGCCTTCCCTTCATCCCGTAGCT-3' (SEQ ID N°9). On procède ensuite, selon le protocole déjà décrit aux points 1 ou 2 précédents, pour la réalisation du plasmide d'expression correspondant.

## 4. Insertion par substitution dans une région ne possédant pas de sites de restriction uniques

L'insertion nécessite au préalable la création des deux sites de restriction Sst I et Xho 1. Pour l'insertion partielle en soi, les deux oligonucléotides suivants ont été synthétisés : 5'-ACATGGTAGAGTAGGTATCTATTTTCGGTATGAAA-3' (SEQ ID N°10) et 5'-TCGATTTTCATACCGAAATAGATACCTACTCTACCATGTAGCT-3' (SEQ ID N°11). Dans ce dernier cas, les résidus R(186) à A(191) sont délétés et substitués par le peptide 11. Les deux plasmides d'expression de ces chimères sont respectivement 1671 et 1667.

Chacun des plasmides, obtenus aux points 1, 2, 3 et 4, est utilisé pour transformer une souche de levure selon le protocole décrit en exemple 4. les protéines correspondantes sont secrétées et purifiées selon les exemples 5 et 6.

5

**EXEMPLE 8****INSERTION DE LA SEQUENCE PEPTIDIQUE IEGR, SUBSTRAT DU FACTEUR Xa**

On procède à l'insertion du peptide contenant la séquence IEGR (SEQ ID N°21), cible du facteur protéasique Xa, qui coupe la prothrombine en thrombine dans la cascade de réactions intervenant dans la coagulation sanguine.

**1.Insertion stricte dans la région 5 ayant un site PvuII unique**

Pour effectuer cette insertion au site Pvu II du gène de l'albumine, il est dans un premier temps créer un vecteur réplcatif dépourvu du site Pvu II, dans lequel est ensuite inséré le gène codant pour la prépro albumine. Les deux oligonucléotides complémentaires suivants, codant pour la séquence IEGR, ont également été synthétisés: 5'-GATCCATAGAAGGTCGACTAG-3'(SEQ ID N°12) et 3'-CTAGGTATCTTCCAGCTGATC-5'(SEQ ID N°13). Ces deux oligonucléotides sont ensuite hybridés puis insérés au niveau du site Pvu II du gène de l'albumine. Les modifications que leur insertion fait apparaître sur les séquences nucléotidiques et peptidiques de l'albumine sont reportées sur la figure 4. La stratégie de clonage est schématisée en figure 5.

Cette construction illustre le cas où des séquences de jonction sont introduites de part et d'autre du peptide d'intérêt.

**2.Insertion dans la région 13.**

La figure 6 décrit une stratégie de clonage d'un peptide dans la région 13.

Dans ce qui suit, les sites de restriction Sst I et Xho 1 ont été créés par mutagenèse dirigée. Dans le cas de la région 13 on a procédé respectivement à une substitution totale et une substitution addition de la séquence IEGR simple ou encadrée de séquences de jonction.



En ce qui concerne la substitution totale, les oligonucléotides utilisés sont les suivants:

5'- CAGAATCGAAGGTAGAGCC-3' (SEQ ID N°14) et

5'- TCGAGGCTCTACCTTCGATCGAGGGTAGCT- 3'(SEQ ID N°15)

- 5        Au niveau protéique, la séquence (187)DEGK (SEQ ID N°22) est substituée par IEGR.

Dans le second cas, les oligonucléotides utilisés sont les suivants:

5'- ACCCTCGATCGAAGGTAGATCTCCA- 3'(SEQ ID N°16)

- 10    5'- TCGATGGAGATCTACCTTCGATCGAGGGTAGCT-3'(SEQ ID N°17)

Au niveau protéique, l'albumine est amputée du résidu R(186) au résidu A(191) et remplacée par la séquence PSIEGRSP (SEQ ID N°23) conduisant donc à une addition de deux résidus.

- 15       Les protéines correspondantes sont secrétées et purifiées suivant les protocoles décrits dans les exemples précédents. le tableau I ci-après rend compte des caractéristiques de ces chimères.

Albumine incorporant:	Taux d'expression ug/l	Rendement purification	Pureté finale
IEGR	150	15	98
PSIEGRPS	110	36	96

TABLEAU I

20

### 3 Activité biologique de la chimère obtenue selon l'exemple 8.1

- 25       La chimère est incubée en présence de facteur Xa bovin dans un rapport enzyme/substrat de 1/10 et dans un tampon Tris 50mM, NaCl 100mM, CaCl<sub>2</sub> 1mM, pH 8.0 pendant 3 heures à 37°C. A l'issue de ce traitement, on réalise une analyse SDS PAGE qui met en évidence le clivage de la chimère caractérisé par la génération d'un fragment d'albumine de masse de l'ordre de 60kD.

30

LISTE DE SEQUENCES

## (1) INFORMATION GENERALE:

- (i) DEPOSANT:
- 5 (A) NOM: RHONE-POULENC RORER S.A.  
(B) RUE: 20, avenue Raymond ARON  
(C) VILLE: ANTONY  
(E) PAYS: FRANCE  
(F) CODE POSTAL: 92165
- 10 (ii) TITRE DE L' INVENTION :Nouveaux polypeptides biologiquement actifs,  
leur preparation et composition pharmaceutique les contenant.
- (iii) NOMBRE DE SEQUENCES: 24
- (iv) FORME LISIBLE PAR ORDINATEUR:
- 15 (A) TYPE DE SUPPORT: Tape  
(B) ORDINATEUR: IBM PC compatible  
(C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS  
(D) LOGICIEL: PatentIn Release #1.0, Version #1.25 (OEB)

## (2) INFORMATION POUR LA SEQ ID NO: 1:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
- 20 (A) LONGUEUR: 13 paires de base  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- 25 (iii) HYPOTHETIQUE: NON  
(iii) ANTI-SENS: NON  
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

GGCCNNNNNG GCC

13

30

## (3) INFORMATION POUR LA SEQ ID NO: 2:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 8 paires de base
- 35 (B) TYPE: acide nucléique  
(C ) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (iii) HYPOTHETIQUE: NON
- (iii) ANTI-SENS: NON
- 40 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

GCGGCCGC

8

## (4) INFORMATION POUR LA SEQ ID NO: 3

- 5 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 30 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire  
10 (ii) TYPE DE MOLECULE: ADNc  
(iii) HYPOTHETIQUE: NON  
(iii) ANTI-SENS: NON  
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

GAAATGCATA AGCTCTTGCC ATTCTCACCG

30

## 15 (5) INFORMATION POUR LA SEQ ID NO: 4:

- 20 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 64 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire  
(ii) TYPE DE MOLECULE: ADNc  
(iii) HYPOTHETIQUE: NON  
(iii) ANTI-SENS: NON  
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°4

25

CCATGGTAGA GTAGGTATCT ATTTCGGTAT GAAACTCCA ACTCTTGTAG AGGTCTCGAG 60  
AAAT 64

## 30 (6) INFORMATION POUR LA SEQ ID NO: 5

- 35 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 68 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire  
(ii) TYPE DE MOLECULE: ADNc  
(iii) HYPOTHETIQUE: NON  
(iii) ANTI-SENS: NON  
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°5

40

CTAGATTCTTCT CGAGACCTCT ACAAGATGTG GAGTTTTCAT ACCGAAATAG ATACCTACTC 60  
TACCATGG 68

45

26

- (7) INFORMATION POUR LA SEQ ID NO: 6:  
    (i) CARACTERISTIQUES DE LA SEQUENCE:  
        (A) LONGUEUR: 34 paires de bases  
        (B) TYPE: acide nucléique  
5        (C) NOMBRE DE BRINS: simple  
        (D) CONFIGURATION: linéaire  
    (ii) TYPE DE MOLECULE: ADNc  
    (iii) HYPOTHETIQUE: NON  
    (iii) ANTI-SENS: NON  
10    (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°6

CCATGGTAGA GTAGGTATCT ATTCGGTAT GAAA

34

- 15 (8) INFORMATION POUR LA SEQ ID NO: 7:  
    (i) CARACTERISTIQUES DE LA SEQUENCE:  
        (A) LONGUEUR: 40 paires de bases  
        (B) TYPE: acide nucléique  
        (C) NOMBRE DE BRINS: simple  
20        (D) CONFIGURATION: linéaire  
    (ii) TYPE DE MOLECULE: ADNc  
    (iii) HYPOTHETIQUE: NON  
    (iii) ANTI-SENS: NON  
    (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°7

25

CTAGTTTCAT ACCGAAATAG ATACCTACTT CTTACCATGG

40

- (9) INFORMATION POUR LA SEQ ID NO: 8:  
30    (i) CARACTERISTIQUES DE LA SEQUENCE:  
        (A) LONGUEUR: 52 paires de bases  
        (B) TYPE: acide nucléique  
        (C) NOMBRE DE BRINS: simple  
        (D) CONFIGURATION: linéaire  
35    (ii) TYPE DE MOLECULE: ADNc  
    (iii) HYPOTHETIQUE: NON  
    (iii) ANTI-SENS: NON  
    (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°8

40 ACGGGATGAA GGAAGGCC ATGCTAGAGT AGGTATCTAT TTCGGTATGA AA

52

- (10) INFORMATION POUR LA SEQ ID NO: 9:  
    (i) CARACTERISTIQUES DE LA SEQUENCE:  
45        (A) LONGUEUR: 61 paires de bases  
        (B) TYPE: acide nucléique  
        (C) NOMBRE DE BRINS: simple  
        (D) CONFIGURATION: linéaire  
    (ii) TYPE DE MOLECULE: ADNc  
50    (iii) HYPOTHETIQUE: NON  
    (iii) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°9

TCGATTTTCAT TACCGAAATA GATACCTACT CTACCATGGG CCTTCCCTTC ATCCCGTAGC 60  
T 61

5

(11) INFORMATION POUR LA SEQ ID NO: 10:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 10 (A) LONGUEUR: 34 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(iii) HYPOTHETIQUE: NON

15 (iii) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°10

ACATGGTAGA GTAGGTATCT ATTCGGTAT GAAA

34

20 (12) INFORMATION POUR LA SEQ ID NO: 11:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 42 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
25 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(iii) HYPOTHETIQUE: NON

(iii) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°11

30

TCGATTTTCAT ACCGAAATAG ATACCTACTC TACCATGTAG CT

42

(13) INFORMATION POUR LA SEQ ID NO: 12:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 35 (A) LONGUEUR: 21 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

40 (iii) HYPOTHETIQUE: NON

(iii) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 12:

GATCCATAGA AGGTCGACTA G

21

45

(14) INFORMATION POUR LA SEQ ID NO: 13:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 21 paires de bases  
(B) TYPE: acide nucléique

28

- (C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire  
(ii) TYPE DE MOLECULE: ADNc  
(iii) HYPOTHETIQUE: NON  
5 (iii) ANTI-SENS: NON  
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°13

CTAGGTATCT TCCAGCTGAT C

21

- 10 (15) INFORMATION POUR LA SEQ ID NO: 14:  
(i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 19 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
15 (D) CONFIGURATION: linéaire  
(ii) TYPE DE MOLECULE: ADNc  
(iii) HYPOTHETIQUE: NON  
(iii) ANTI-SENS: NON  
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°14

20

CAGAATCGAA GGTAGAGCC

19

- (16) INFORMATION POUR LA SEQ ID NO: 15:  
(i) CARACTERISTIQUES DE LA SEQUENCE:  
25 (A) LONGUEUR: 30 paires de bases  
(B) TYPE: acide nucléique  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire  
(ii) TYPE DE MOLECULE: ADNc  
30 (iii) HYPOTHETIQUE: NON  
(iii) ANTI-SENS: NON  
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°15

TCGAGGCTCT ACCTTCGATC GAGGGTAGCT

30

- 35 (17) INFORMATION POUR LA SEQ ID NO: 16:  
(i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 24 paires de bases  
(B) TYPE: acide nucléique  
40 (C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire  
(ii) TYPE DE MOLECULE: ADNc  
(iii) HYPOTHETIQUE: NON  
(iii) ANTI-SENS: NON  
45 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°16

ACCCTCGATC GAAGGTAGAT CTCCA

24

- 50 (18) INFORMATION POUR LA SEQ ID NO: 17:  
(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 33 paires de bases  
 (B) TYPE: acide nucléique  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire
- 5 (ii) TYPE DE MOLECULE: ADNc  
 (iii) HYPOTHETIQUE: NON  
 (iii) ANTI-SENS: NON  
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°17
- 10 TCGATGGAGA TCTACCTTCG ATCGAGGGTA GCT 33
- (19) INFORMATION POUR LA SEQ ID NO: 18:  
 (i) CARACTERISTIQUES DE LA SEQUENCE:  
 (A) LONGUEUR: 31 paires de bases  
 15 (B) TYPE: acide nucléique  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire  
 (ii) TYPE DE MOLECULE: ADNc  
 (iii) HYPOTHETIQUE: NON  
 20 (iii) ANTI-SENS: NON  
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°18
- GTCCCGGATG GAGCGCGTAC TTAGAGAGAA T 31
- 25 (20) INFORMATION POUR LA SEQ ID NO: 19:  
 (i) CARACTERISTIQUES DE LA SEQUENCE:  
 (A) LONGUEUR: 39 paires de bases  
 (B) TYPE: acide nucléique  
 (C) NOMBRE DE BRINS: simple  
 30 (D) CONFIGURATION: linéaire  
 (ii) TYPE DE MOLECULE: ADNc  
 (iii) HYPOTHETIQUE: NON  
 (iii) ANTI-SENS: NON  
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°19
- 35 TCGACAGGGC CTACCTCGCG CATGAATCTC TCTTAAGCT 39
- (21) INFORMATION POUR LA SEQ ID NO: 20:  
 (i) CARACTERISTIQUES DE LA SEQUENCE:  
 40 (A) LONGUEUR: 11 acides aminés  
 (B) TYPE: acide aminé  
 (D) CONFIGURATION: linéaire  
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 20
- 45 His Gly Arg Val Gly Ile Tyr Phe Gly Met Lys

30

(22) INFORMATION POUR LA SEQ ID NO: 21:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 4 acides aminés

(B) TYPE: acide aminé

5 (D) CONFIGURATION: linéaire

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 21

Ile Glu Gly Arg

10 (23) INFORMATION POUR LA SEQ ID NO: 22:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 4 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

15 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 22

Asp Glu Gly Lys

(24) INFORMATION POUR LA SEQ ID NO: 23:

20 (i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 8 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 23

25

Pro Ser Ile Glu Gly Arg Ser Pro

5

30 (25) INFORMATION POUR LA SEQ ID NO: 24:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 9 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

35 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 24

Arg Met Glu Arg Val Leu Arg Glu Asn

5



### REVENDEICATIONS

1. Polypeptide recombinant comportant au moins une partie active dérivée d'un polypeptide, naturel ou synthétique, biologiquement actif, génétiquement insérée dans une albumine ou un de ses variants ou dérivés.
- 5 2. Polypeptide selon la revendication 1 caractérisé en ce que le polypeptide biologiquement actif possède une activité thérapeutique et est d'origine humaine.
3. Polypeptide selon la revendication 2 caractérisé en ce que le polypeptide ayant une activité thérapeutique est choisi parmi tout ou partie des enzymes, des inhibiteurs d'enzymes, des antigènes, des anticorps, des hormones, des récepteurs,  
10 des facteurs de la coagulation, des interférons, des cytokines, des facteurs de croissance et/ou de différenciation, des facteurs impliqués dans la génèse/résorption des tissus osseux, des facteurs chimiotactiques, des facteurs de motilité ou de migration cellulaire, des facteurs cytostatiques, des facteurs bactéricides ou antifongiques, ou des molécules adhésives plasmatiques, interstitielles ou des  
15 matrices extracellulaires.
4. Polypeptide selon l'une des revendications 1 à 3 caractérisé en ce que le polypeptide ayant une activité thérapeutique est choisi parmi toute séquence peptidique antagoniste ou agoniste d'interactions moléculaires et/ou cellulaires impliquées dans les pathologies des compartiments circulatoires et interstitiels.
- 20 5. Polypeptide selon l'une des revendications 1 à 4 caractérisé en ce que la partie active présente une structure choisie parmi :
  - (a) la structure peptidique entière ou,
  - (b) un fragment de (a) ou une structure dérivée de (a) par modification structurale (mutation, substitution, addition et/ou délétion d'un ou plusieurs résidus) et  
25 conservant une activité thérapeutique.
6. Polypeptide selon l'une des revendications 1 à 5 caractérisé en ce que la partie active est insérée strictement à l'intérieur de l'albumine ou entourée de séquence de jonctions.

7. Polypeptide selon l'une des revendications 1 à 6 caractérisé en ce que la partie active est insérée de préférence au niveau des régions de l'albumine présumées former des régions exposées à la surface de la molécule.
8. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la  
5 partie active est insérée au niveau de la région 5 s'étendant du résidu 57 à 62 de l'albumine.
9. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la partie active est insérée au niveau de la région 8 s'étendant du résidu 103 à 120 de l'albumine.
10. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la  
10 partie active est insérée au niveau de la région 13 s'étendant du résidu 178 à 200 de l'albumine.
11. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la partie active est insérée au niveau de la région du résidu 415 au résidu 425 délimitée  
15 par les hélices h2 et h3 du domaine III dans l'albumine.
12. Polypeptide selon l'une des revendications 1 à 12 caractérisé en ce que la partie active y est insérée sous forme unique ou multiple.
13. Polypeptide selon la revendication 12 caractérisé en ce que la partie active est répéter plusieurs fois au même endroit et/ou dans des régions différentes de  
20 l'albumine.
14. Polypeptide selon l'une des revendications 1 à 13 caractérisé en ce que les parties actives insérées sont de nature différente.
15. Variant d'une séquence nucléotidique codant pour l'albumine ou un de ses variants ou dérivés intégrant au moins un site de restriction unique non naturel.
16. Séquence nucléotidique codant pour un polypeptide selon l'une quelconque  
25 des revendications 1 à 14.
17. Séquence nucléotidique selon la revendication 16 caractérisée en ce qu'elle comprend une séquence "leader" permettant la sécrétion du polypeptide exprimé.

18. Cassette d'expression comprenant une séquence nucléotidique selon l'une des revendications 16 ou 17 sous le contrôle d'une région d'initiation de la transcription et éventuellement d'une région de terminaison de la transcription.
19. Plasmide autoréplcatif comportant une cassette d'expression selon la  
5 revendication 18.
20. Cellule recombinante eucaryote ou procaryote dans laquelle a été inséré une séquence nucléotidique selon l'une des revendications 16 ou 17 ou une cassette d'expression selon la revendication 18 ou un plasmide selon la revendication 19.
21. Cellule recombinante selon la revendication 20 caractérisée en ce qu'il s'agit  
10 d'une levure, d'une cellule animale, d'un champignon ou d'une bactérie.
22. Cellule recombinante selon la revendication 21 caractérisée en ce qu'il s'agit d'une levure.
23. Cellule recombinante selon la revendication 22 caractérisée en ce qu'il s'agit d'une levure du genre Saccharomyces ou Kluyveromyces.
- 15 24. Procédé de préparation d'un polypeptide tel que défini dans l'une des revendications 1 à 14 caractérisé en ce que l'on cultive une cellule recombinante selon l'une des revendications 20 à 23 dans des conditions d'expression, et on récupère le polypeptide produit.
25. Composition pharmaceutique comprenant un ou plusieurs polypeptides selon  
20 l'une quelconque des revendications 1 à 14.
26. Composition pharmaceutique comprenant une séquence nucléotidique selon l'une quelconque des revendications 16 à 17 utilisable en thérapie génique.

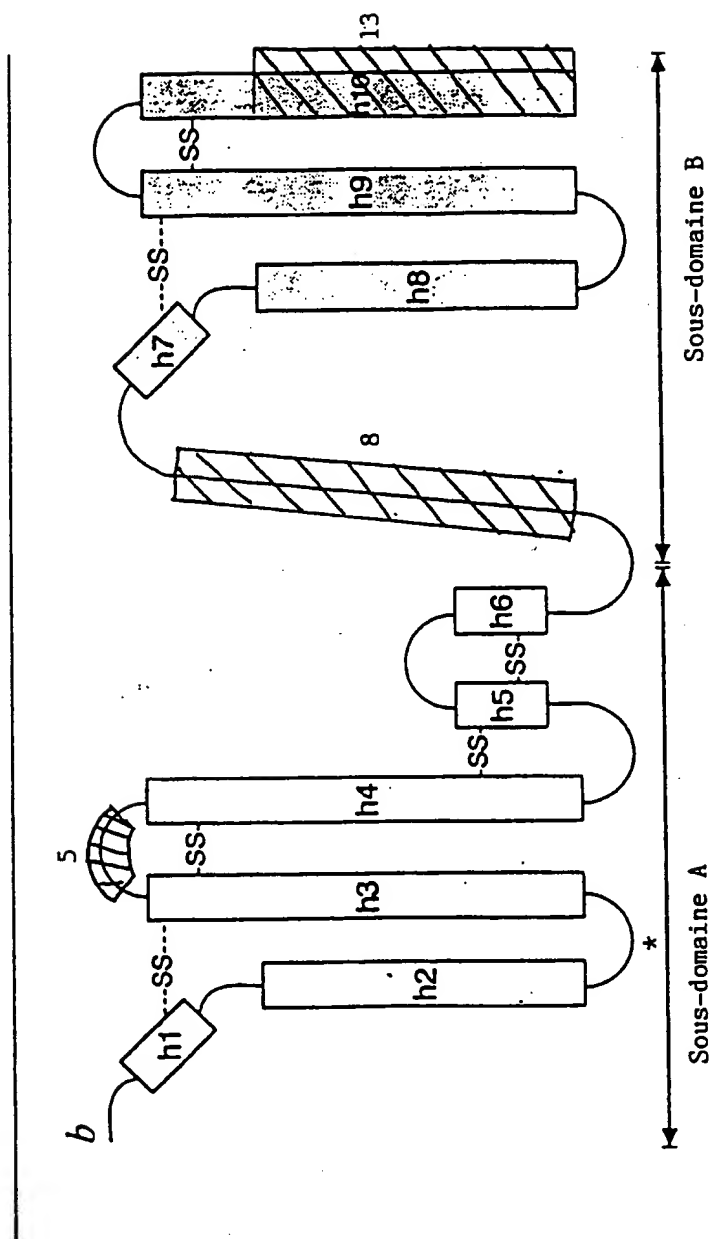


Figure 1

2/6

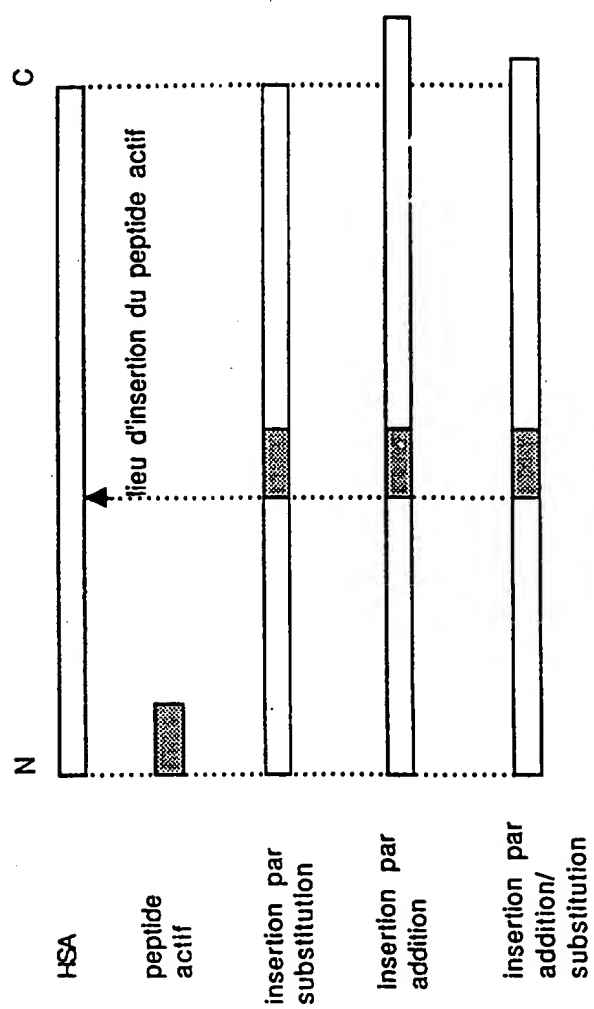


Figure 2

3/6

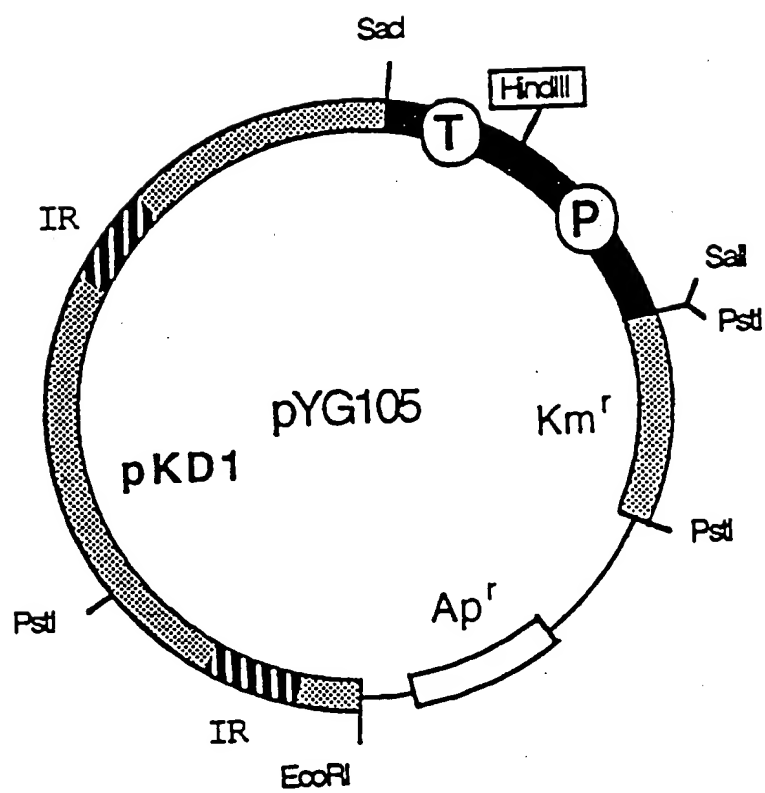


Figure 3



5/6

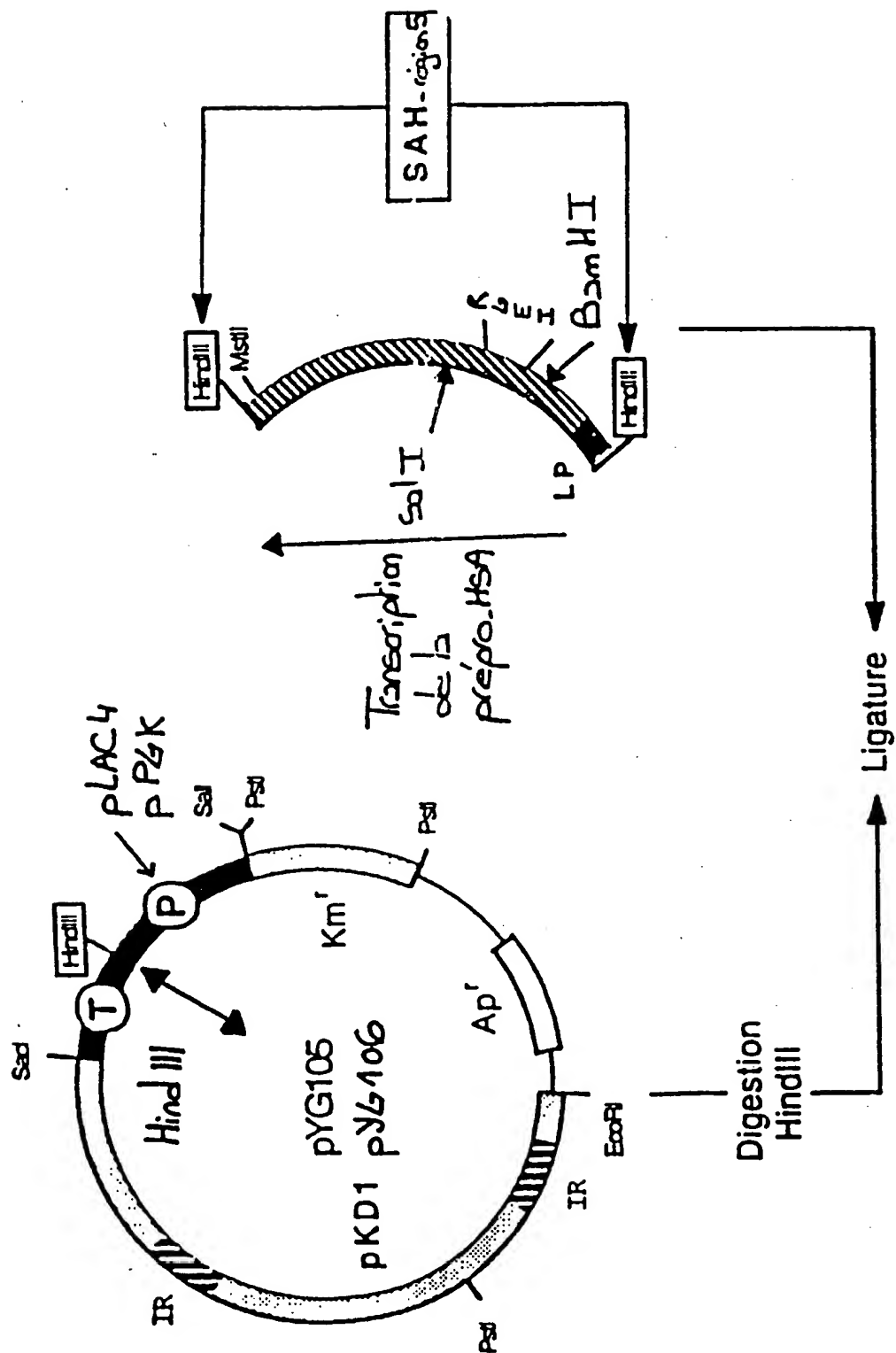


Figure 5



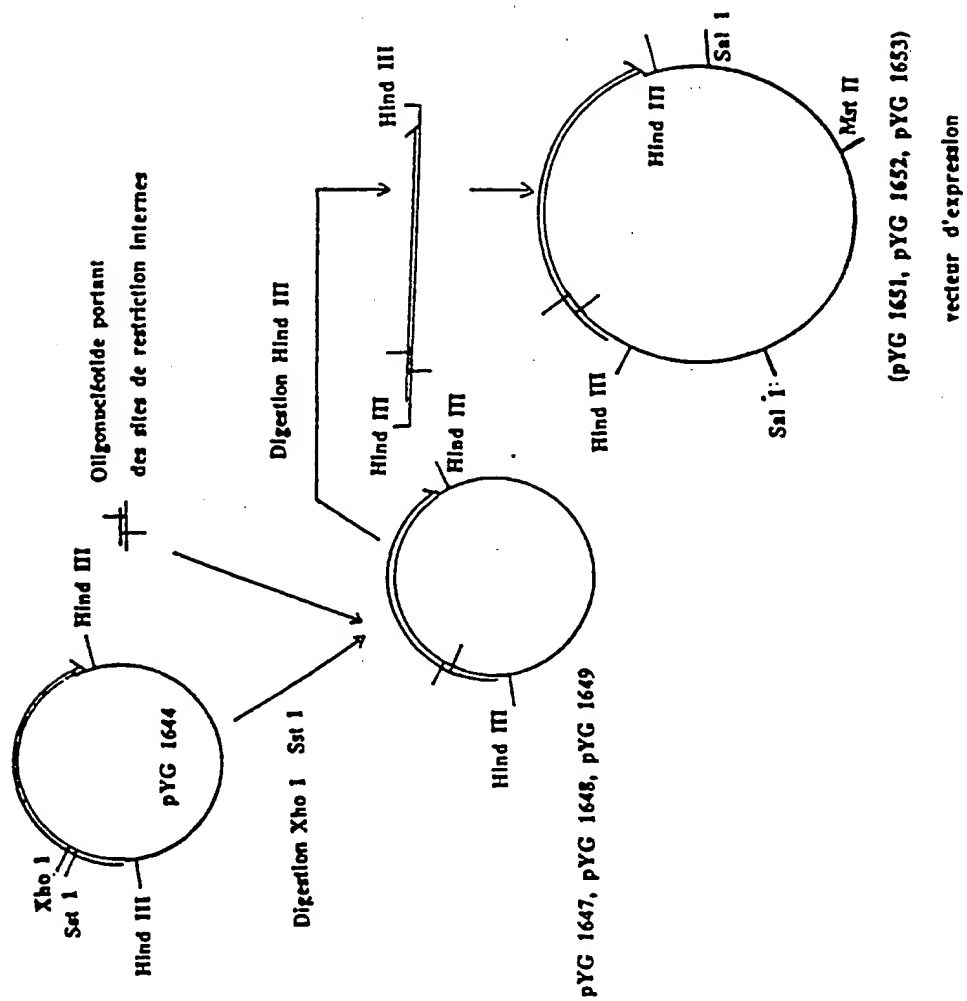


Figure 6

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/FR 95/00520

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12N15/62 C12N15/14 C12N15/12 C12N15/60 C12N15/80  
C12N15/81 C12N1/19 C12N1/15 A61K38/38 A61K38/17  
C07K14/765 //C12N9/88, C12N1/21, C12N5/10,

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO-A-93 15199 (RHONE-POULENC-RORER S.A) 5 August 1993 see the whole document ---	1-26
A	BIOTECHNOLOGY, vol. 7, September 1989 pages 929-932, J. VANDEKERCKHOVE ET AL. 'Enkephalins produced in transgenic plants using modified 2S seed storage proteins' see the whole document ---	1-3, 5, 6, 12, 16-20, 25
A	EP-A-0 399 666 (DELTA BIOTECHNOLOGY LIMITED) 28 November 1990 * see the whole document, in particular examples 2 and 3 * --- -/-	1-5, 7, 10, 12, 15-26

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to undermine the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*A\* document member of the same patent family

Date of the actual completion of the international search

14 August 1995

Date of mailing of the international search report

23.08.95

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+ 31-70) 340-3016

Authorized officer

Gac, G

## INTERNATIONAL SEARCH REPORT

Inter national Application No

PCT/FR 95/00520

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO-A-93 15200 (RHONE-POULENC-RORER S.A.) 5 August 1993 see the whole document ---	1-26
A	WO-A-93 15211 (RHONE-POULENC RORER S.A.) 5 August 1993 see the whole document ---	1-26
A	J. MOL. BIOL., vol. 184, no. 4, 20 August 1985 pages 547-564, JUBIER-MAURIN ET AL. 'Comparative study of the L1 family in the genus Mus : possible role of retroposition and conversion events in its concerted evolution' * pages 547, 552 * * page 555 right column * -----	1,5,6, 12,16

# INTERNATIONAL SEARCH REPORT

Inter-  
national Application No  
PCT/FR 95/00520

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
WO-A-9315199	05-08-93	FR-A-	2686899	06-08-93
		EP-A-	0624195	17-11-94
		FI-A-	943563	29-07-94
		JP-T-	7503368	13-04-95
-----				
EP-A-0399666	28-11-90	DE-T-	69002395	25-11-93
		EP-A-	0407008	09-01-91
		ES-T-	2060033	16-11-94
		WO-A-	9013306	15-11-90
-----				
WO-A-9315200	05-08-93	FR-A-	2686901	06-08-93
		CA-A-	2126092	05-08-93
		EP-A-	0625199	23-11-94
		FI-A-	943565	29-07-94
		JP-T-	7503369	13-04-95
		NO-A-	942840	29-09-94
-----				
WO-A-9315211	05-08-93	FR-A-	2686900	06-08-93
		CA-A-	2125979	05-08-93
		EP-A-	0624200	17-11-94
		FI-A-	943564	29-07-94
		JP-T-	7503844	27-04-95
		NO-A-	942858	01-08-94
-----				

## RAPPORT DE RECHERCHE INTERNATIONALE

 Dem: Internationale No  
 PCT/FR 95/00520

A. CLASSEMENT DE L'OBJET DE LA DEMANDE  
 CIB 6 C12N15/62 C12N15/14 C12N15/12 C12N15/60 C12N15/80  
 C12N15/81 C12N1/19 C12N1/15 A61K38/38 A61K38/17  
 C07K14/765 //C12N9/88, C12N1/21, C12N5/10,

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

## B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 C07K C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)

## C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO-A-93 15199 (RHONE-POULENC-RORER S.A) 5 Août 1993 voir le document en entier ---	1-26
A	BIOTECHNOLOGY, vol. 7, Septembre 1989 pages 929-932, J. VANDEKERCKHOVE ET AL. 'Enkephalins produced in transgenic plants using modified 2S seed storage proteins' voir le document en entier ---	1-3, 5, 6, 12, 16-20, 25
A	EP-A-0 399 666 (DELTA BIOTECHNOLOGY LIMITED) 28 Novembre 1990  *le document en entier, surtout exemples 2 et 3 * --- -/--	1-5, 7, 10, 12, 15-26

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

## \* Catégories spéciales de documents cités:

- \*A\* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- \*E\* document antérieur, mais publié à la date de dépôt international ou après cette date
- \*L\* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- \*O\* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- \*P\* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

\*T\* document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

\*X\* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

\*Y\* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

\*Z\* document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

14 Août 1995

Date d'expédition du présent rapport de recherche internationale

23.08.95

Nom et adresse postale de l'administration chargée de la recherche internationale  
 Office Européen des Brevets, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tél. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax (+31-70) 340-3016

Fonctionnaire autorisé

Gac, G

# RAPPORT DE RECHERCHE INTERNATIONALE

Dem Internationale No  
PCT/FR 95/00520

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO-A-93 15200 (RHONE-POULENC-RORER S.A.) 5 Août 1993 voir le document en entier ---	1-26
A	WO-A-93 15211 (RHONE-POULENC RORER S.A.) 5 Août 1993 voir le document en entier ---	1-26
A	J. MOL. BIOL., vol. 184, no. 4, 20 Août 1985 pages 547-564, JUBIER-MAURIN ET AL. 'Comparative study of the L1 family in the genus Mus : possible role of retroposition and conversion events in its concerted evolution' * pages 547, 552 * * page 555 colonne de droite * -----	1,5,6, 12,16

# RAPPORT DE RECHERCHE INTERNATIONALE

Dem Internationale No  
PCT/FR 95/00520

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)		Date de publication
WO-A-9315199	05-08-93	FR-A-	2686899	06-08-93
		EP-A-	0624195	17-11-94
		FI-A-	943563	29-07-94
		JP-T-	7503368	13-04-95
-----				
EP-A-0399666	28-11-90	DE-T-	69002395	25-11-93
		EP-A-	0407008	09-01-91
		ES-T-	2060033	16-11-94
		WO-A-	9013306	15-11-90
-----				
WO-A-9315200	05-08-93	FR-A-	2686901	06-08-93
		CA-A-	2126092	05-08-93
		EP-A-	0625199	23-11-94
		FI-A-	943565	29-07-94
		JP-T-	7503369	13-04-95
		NO-A-	942840	29-09-94
-----				
WO-A-9315211	05-08-93	FR-A-	2686900	06-08-93
		CA-A-	2125979	05-08-93
		EP-A-	0624200	17-11-94
		FI-A-	943564	29-07-94
		JP-T-	7503844	27-04-95
		NO-A-	942858	01-08-94
-----				